https://doi.org/10.15407/fd2024.04.130
UDC: 167.7

Anna LAKTIONOVA,
Dr., Prof., Chair of Philosophy of Science and Technology,
RWTH, Aachen (Germany);
Professor at Department of Theoretical and Practical Philosophy,
Faculty of Philosophy, Taras Shevchenko National University of Kyiv,
60, Volodymyrska St., Kyiv (Ukraine), 01601
laktionovaanna@yahoo.com
https://orcid.org/0000-0001-8236-7217

HUMAN-MACHINE INTERACTIONS: ALIGNING, ADAPTING, BEING AN AGENT[1]

In the paper, the touchstone points of the project "Towards an agency-based philosophy of (advanced) technology" are outlined. The main plot of this elaboration concerns human-machine interactions and appropriate interpretation of reciprocal aligning, adapting within involved into such interactions agents; as well as the status as such of being an agent. Into the theoretical and historical background of the project such spheres as Philosophy of Science, Philosophy of Technology, Philosophy of Engineering and Design Technological Actions, STS (Science and Technology Studies), Applied Ethics etc. could be invited. To treat agency as technology, reestablish the role of agency in technology is the most ambitious goal of the project: 'activity as technology' focuses on activities through technologies. The terms «agency» and «activity» are used in this paper synonymously with the basic Aristotelian meaning of agent's potential capacity to act. The proposed by the author theory of action and agency (and cor-

[1] Acknowledgements: This paper was prepared within the DFG Grant, project "Towards an Agency based Philosophy of (advanced) Technology", the number for publications is 536935318. I am very thankful to Dr., Prof. Gabriele Gramelsberger, chair for Theory of Science and Technology at the Faculty of Arts and Humanities, RWTH, Aachen University, for general and personal support of this project.

Citation: Laktionova, A. (2024). Human-machine interactions: aligning, adapting, being an agent. *Філософська думка*, 4, 130—142. https://doi.org/10.15407/fd2024.04.130

respondently defended in 2016 dissertation of Doctor of Philosophical Science) is to be applied into philosophical reflections about various problems of dealing with currently continually appearing flourishing fields of artificial intelligence, machine learning (including deep learning methods and simulation methods), inventing computers and codes, numerically controlled machines and robots, computer-chip equipped devices, smart objects etc.

The described in the article proposed research project on human-machine interaction can get the empirical materials from collaboration with at least the following "labs" at RWTH Aachen University: the Institute of Industrial Engineering and Ergonomics (https://www.iaw.rwth-aachen.de/cms/~ieplw/IAW/lidx/1/); the Center for Construction Robotics (https://construction-robotics.de/en/); the Chair Individual and Technology (https://www.itec.rwth-aachen.de/cms/~sjbp/ITEC/?lidx=1).

Keywords: *agency, (inter)actions, agent, machine, aligning, adapting, artificial intelligence.*

The approach to view technology as activity had been quite widespread for years. In "Thinking Through Technology: The Path Between Engineering and Philosophy" Carl Mitcham investigated the various traditions of philosophy of technology and distinguished, in particular, two distinct views on technology (Mitcham, 1994): the engineering philosophy of technology and the humanities philosophy of technology, the first explaining the nature of technology from inside (e.g.: Kapp, 2018; Dessauer, 1972), the latter trying to understanding technology from a genuine philosophical perspective and, thus, establishing the primacy of humanities over technology (e.g.: Heidegger, 1993; Ellul, 1964; Mumford, 1967—70). However, for Mitcham as well as his references from both sides, technology is the making and using of artifacts usually focusing on either 'technology as object' or 'technology as knowledge.' Instead of following these traditions, Mitcham introduced the less common approach of 'technology as activity.' "Technology as activity is that pivotal event in which knowledge and volition unite to bring artifacts into existence or to use them; it is likewise the occasion for artifacts themselves to influence the mind and will" (Mitcham, 1994: p. 109). The 'technology as activity' approach interlinks artifacts with human actions (e.g., crafting, inventing, designing, operating, maintaining, working). While the activities to bring artifacts into existence or to use them put humans in the role of agents, the occasion for artifacts themselves to influence the mind and will reminds mysterious. Technology as activity was linked by Mitcham to the various human behaviors, but not to machinic ones.

One specific human behavior, already mentioned in (Mitcham, 1994), is engineering covering many other above-mentioned actions. The emerging philosophy of engineering explores the nature of this multiplex action of engineering, what engineers are doing, and how these engineering actions influence society and lifeworld (Mitcham, 2019; Dias, 2019). Aiming at initiating a critical reflection among engineers and non-engineers on ethical topics and a better engineering self-understanding (Mitcham, 2019), this interdisciplinary dialog is increasingly becoming a current practice in technology development. Responsible Research and Innovation (e.g.: Schomberg, 2013; EU, 2013), ethics by design (e.g.: Dignum et al., 2018) and the IEEE Standard Model Process for Addressing Ethical Concerns during System Design (IEEE, 2021), respectively, have created the new field of applied ethics of technology.

Nevertheless, machine agency becomes central, when technology's activities are increasingly steered and controlled by software and the radius of machine agency becomes ever larger and more autonomous. Computers, but also numerically controlled machines and robots, computer-chip equipped devices and smart objects not only have become widespread technical artifacts, but have established a new type of technology. My main hypothesis for this paper is that technology as activity has turned into 'activity as technology', and the necessary preconditions are sensors, actuators, algorithms and software. For instance, behavioral robotics, the basis of today's robot design, introduced by Rodney Brooks, puts activity in lifeworld environment first instead of programmed behavior: "We have developed a computational architecture known as the subsumption architecture. It enables us to tightly connect perception to action, embedding robots concretely in the world" (Brooks, 1990: p. 5). Today, behavioral robotics is increasingly linked with machine learning (ML) technologies constituting the new paradigm of adaptive robotics (Gramelsberger, 2023: p. 210 et seq.).

Treating activity as technology opens a fascinatingly interesting role of agency in technology. Generally speaking, agency is the capacity to act in a surrounding environment. From this capacity many forms of concrete activities and purposeful actions follow such as walking, listening, seeing, producing, etc. The philosophy of action explores the motivation of agency. As agency is regarded as human agency, reasons and intentions have been identified as the dominant concepts for motivating actions (e.g.: Davidson, 1963; Goldman, 1970). This standard account to agency identifies actions with events, if and only if it is an action under some description. The same action can be intentional for some instances (descriptions) and unintentional for others. Harry G. Frankfurt distinguished between the agency of persons and other agents, in which persons reflect on their motivation to act (Frankfurt, 1971, 1978). However, other agents are usually not the subjects of the philosophy of action due to its intrinsic characterization of agency usually assignable to humans only. Or, to put it differently: Do non-human agents reflect, follow reasons and intentions? Or can, at least, a form of "minimal agency" be assigned to non-humans (Barandiaran et al., 2009)? Must such non-humans be organisms, or can they be machines? In science and technology studies (STS) and sociology of technology the actor-network theory (ANT) assigns equal agency to humans and objects within a constantly shifting networks of relationships (e.g.: Law, 1992; Latour, 2005). Because of its strict empirical approach and the presumption that nothing exists outside those relationships and that agency is a metaphysical concept, ANT only allows to describe actions rather than to conceptually analyze and explain them. Thus, for the proposed philosophical perspective ANT is not an appropriate methodology.

Further, beside classical forms like intentional agency, philosophy has to (and already did) consider other forms of agency such as mental, epistemic, shared, collective, relational and artificial agency (for an overview see: Schlosser, 2019; Roth, 2017). In particular, artificial agency is of interest for an 'activity as technology' approach, but also shared agency (Searl, 1990). Following Daniel Dennett, intentionality can be assigned to machines in some cases for a "system whose behavior is reliably and voluminously predictable via the intentional strategy" like for chess computers (see: Dennett,

1971, 1984). However, in case of machines, the subsumption architecture introduced by Brooks in "Elephants don't play chess" eliminates any rationality and intentionality opposing cognitive AI symbolically programmed into machines: "The traditional [AI] methodology bases its decomposition of intelligence into functional information processing modules whose combinations provide overall system behavior. The new methodology bases its decomposition of intelligence into individual behavior generating modules, whose coexistence and co-operation let more complex behaviors emerge" (Brooks, 1990: p. 3). In this regard, Brooks' subsumption architecture is a good case for "minimal agency" based solely on adaptive regulation through linking with an environment (Barandiaran et al., 2009), but also for the debates on embodied mind and enactivism in the philosophy of mind and cognitive science discussing agency without representational mental states (e.g.: Gallagher, 2005, 2020; Shapiro, 2007, 2013, 2019).

However, today we experience the shift to adaptive robotics interlinking behavioral robotics with machine learning (ML) technologies. Therefore, the question arises: Does connecting perception to action in lifeworld-embedded machine bodies and nowadays combining it with ML algorithms establish more than "minimal agency" for machines? While Brookes' paradigm shift introduced a "minimal agency" model (questioned in: Schlosser, 2018), the ML algorithms enable machinic versions of perception, concept building, imitation learning, reflections and evaluations of the environmental states; hence (re)presentational states.

As outlined above, the various strands of philosophy of action address aspects which can be applied to develop a model of machine agency beyond minimal agency, in particular, by developing Dennett's intentional stance further. As Gabriele Gramelsberger argues in "Philosophie des Digitalen zur Einführung", today's advanced technologies cannot be understood in their lifeworld encounters with us without intentional assignments by us to technology (Gramelsberger, 2023: p. 160 et seq.). Using such advanced technologies requires an alike advanced literacy of machinic intentions, simply because they are programmed into these technologies or have been formed during machine learning setups. Of course, intentions refer here to "minimal intentions," i.e., algorithmic decisions informed by statistically activated threshold-steered behavior. And, as we have experienced over the past years, these machinic minimal intentions can go wrong (machine biases (e.g., false negatives and false positives) and adversarial failures). These biases and failures have been intensively discussed, among other aspects, in the fields of data, machine, and robot ethics (e.g.: Floridi & Sanders, 2002; Capurro & Nagenborg, 2009; Floridi, 2013; Leonelli, 2016; Misselhorn, 2019; Kaminski, 2020; Nyholm, 2023) and have inspired programs like ethics by design (e.g.: Dignum et al., 2018) and *Ethics Guidelines for Trustworthy AI* (EU, 2019). However, beside these ethical aspects a better understanding of the nature and types of the rapidly evolving field of machine agency is required. Philosophy of action offers great opportunities to be developed further for machine agency; and can be directed towards an agency-based philosophy of technology.

During more than 15 years my work focusses on philosophy of action and agency. Already in my monograph *Philosophy of Action* (Laktionova, 2016) and dissertation I outlined new perspectives on actions for theoretical and practical philosophical disciplines.

This included "Collective and Individual Actions" (Laktionova, 2015), reflections on "Meta-Ontological Discourse on Being and Activity" (Laktionova, 2016a) and "Contemporary Philosophy of Action (Practical Philosophy) in Analytic Tradition" (Laktionova, 2016b). Cornerstones of my *Philosophy of Action* are the reciprocally additive character of the relation between 'is' (facts) and 'ought' (norms), the connection between 'being' and 'human agency', validity of personal identity from inside of the experience of free agents, the concept of perceptive activity and experience etc. Different to Donald Davidson (1963), I argue that actions need agents, while events don't. Core for me is the principle of performativity, which I introduced from Philosophy of Language, especially from Ludwig Wittgenstein's concept of "Sprachspiele" (Wittgenstein, 1953) and John L. Austin's Speech Act Theory (Austin, 1962), to philosophy of agency (Laktionova, 2016). I have continually developed the principle of performativity further into a philosophical theory of "Performativity of Actions" (Laktionova, 2022). By performativity I mean the accomplishment of what is being performed by action (realization). Every action is self-demonstrating and self-justifying by being performative, but leaves room for its interpretation, if it initializes a chain of other actions (e.g., picking up a pen can lead to all kind of follow-up actions: writing something, lending it to someone else, packing it into a pencil case, etc.). In particular, Jennifer Hornsby's classical account of "Acting and Trying to Act" (Hornsby, 1980, and her later works in: Hornsby, 2010, 2012) inspires my theory of performativity of agency, because Hornsby's concept of 'trying to act,' if successful (accomplished), is always within the action, not external to it (Laktionova, 2022). Thus, performativity coincides with the action, while intention (first) is always external to the action (second). Because there is no need to relate intentions and attempts (attempts become usually only conscious in the moment they fail), intention to act can be treated in terms of attempts. Thus, a theory of performativity of agency is not necessarily restricted to humans.

From these considerations results my interest in applying my approach of philosophy of actions and agency to philosophy of technology (Laktionova, 2023, 2023a). Based on this background, developing an agency-based philosophy of (advanced) technology becomes possible. The task of interactively elaborating a novel agency-based philosophy of (advanced) technology is related to such prominent topics in current philosophy as general philosophy of the digital, robot technology and enactivism, philosophy of AI, AI cultures of research, lifelikeness of technology, reflections of methods for studying advanced technologies, autonomy of agents etc.; could involve opportunities of collaboration with scientists, scientific laboratories, groups etc. to directly observe the work of researchers from robotics, AI, and other advanced technology areas where various computational studies like code studies and developing of software tools make the interlinkage between technology and agency evident. The clarification of very touchstone concepts for human — non-human interactions, such as alignment, adaptation etc., could appear.

Developing an agency-based philosophy of (advanced) technology from the perspective of philosophy of action and agency could be based on three presumptions:

1. Activity as technology: While studies from science and technology studies (STS) describe cases of technology as activity, applied ethics evaluates and regulates

technology as activity, and philosophy of science investigates the methods of technology as activity, I aim at a genuine philosophical analysis of 'activity as technology.' This inversion is crucial for an agency-based philosophy of (advanced) technology. While 'technology as activity' highlights activities *around* technologies such as crafting, inventing, designing, operating, maintaining, engineering, ethically regulating, social interacting, 'activity as technology' focus on activities *through* technologies, i.e. performed by technology and as machine agency, respectively.

2. Intrinsic normativity of technology: While science is dealing with the mode of "is" (what there is, i.e. facts), technology is dealing with the mode of "ought" (what ought to be, i.e. purposes, norms). Of course, technology in its existence is a fact (for us) and generates its own techno-phenomenological sphere of facts (for us) (Bachelard, 1985; Karafyllis, 2007; Nordmann, 2006). But, for technology, on the one hand, "is" ought to follow from its intended design and purposes (for Dennett's design stance see: Dennett, 1971, 1984), on the other hand, facts, opposite to events, are not processes and a fact cannot be unknown (then it is not a fact). Or, put it differently: Technology generates facts (for us), but lacks knowledge about these facts.

3. Minimal intentionality/intentional agency: From (2) follows that machines can have intentionality/intentional agency, but only a minimal version of it, i.e. adaptive regulation through linking with an environment. Intentions are only ascribable to machines in two ways: on the one hand attributed by design and coding, if formally operationalizable, on the other hand by assigning intentionality to advanced technologies by users (Dennett's intentional stance). Both ways remain to be problematic.

Presumptions (2) and (3) mark the challenges for an agency-based philosophy of technology as advanced technologies realize machinic versions of perception, concept building, reflections and evaluations of the environmental states (even of their own states created by the technology as facts); hence advanced technologies can probably develop (re)presentational states. However, without the presumptions of intrinsic normativity and minimal intentionality/intentional agency of technology one is forced to concede that advanced technologies have a sense of self-reflection, consciousness, and autonomous intelligence, as it is claimed in transhumanistic concepts of AI (e.g.: Bostrom, 2013). The latter is not acceptable. Thus, the real challenge for an agency-based philosophy of technology is to develop a philosophical framework that allows to analyze, understand, and philosophically interpret the ongoing technological transformations without such an acknowledgment.

Within developing an agency-based philosophy of (advanced) technology and taking the above-mentioned challenges into account the main research points look like:

A. Developing a philosophical framework that allows to analyze the ongoing technological transformations towards advanced technologies via describing the notion of advanced technologies as 'activity as technology' (case of adaptive robots); consequently, classifying various types, levels, and scales of machinic actions.

B. Understanding (advanced) technologies from perspective of philosophy of actions and agency via locating the conceptual entry points for the philosophy of action and agency to be applied to (advanced) technologies.

C. Pphilosophically interpreting (advanced) technologies via outlining an agency-based philosophy of (advanced) technologies

Based on Mitcham's classification of four types of technologies (Mitcham, 1994) the characteristic aspects of advanced technologies as 'activity as technology' can be established and illustrated by the case of adaptive robots. For adaptive robots, Brooks' assumption architecture (Brooks,1990) has been enhanced with better sensors and actuators, but in particular with machine learning (ML) methods. Adaptive robots use all kind of ML algorithms, for instance reinforcement learning for imitation learning, unsupervised learning for self-steered development, multi-agent learning for collaborative behaviour, deep learning methods for improving perception and object recognition, and many more ML and simulation methods (Nolfi, 2022). Thus, adaptive robots can be seen as paradigmatic entities for advanced technologies, but in particular for complex machine agency.

Investigation of adaptive robots as paradigmatic entities of machine agency is twofold. First, it is worth to analyse the various concepts of adaptive robotics presented in current literature by focussing on the following sources: Scientific journals like *Frontiers in Robotics and AI*; *Journal of Intelligent & Robotic Systems*; *Robotics Engineering Research*; *arxiv.org*.; Business technology magazines and manufacturers websites like *roboticsbusinessreview.com*; *uscybersecurity.net*; *wired.com*; *infotech.com*; *roboticstomorrow.com*.; Policy reports and whitepapers on robotics and AI like the European Parliament (2017) *Resolution Civil Law Rules on Robotics*.

Such literature could provide better understanding of how the designs of adaptive robots, and in particular agency, is conceived. For instance, as Shiquan Wang outlined in *roboticsbusinessreview.com*: "A robot also needs to know how to utilize force control with the integration of other information. This leads us to the concept of hierarchical intelligence, which naturally applies to people's daily activity" (see: Wang, 2019).

Additionally, if researchers and policy makers identify machinic actions as problematic there remains a question why they mark them as problematic. Do we need ethical regulations for trustworthy adaptive robotics? Can an agency-based philosophy of advanced technology offer support here?

Second, the research on human-robot interaction can be observed at the following three labs at RWTH Aachen University (where I had been a fellow (at Käte Hamburger Kolleg "Cultures of Research") for 2 years): the RWTH Institute of Industrial Engineering and Ergonomics is investigating adaptive human-robot collaboration for the ergonomic design of work (Prof. Dr. Verena Nitsch); the RWTH Center for Construction Robotics is working on adaptive robot control and new parametric workflows (Prof. Dr. Sigrid Brell-Çokcan); the RWTH Chair Individual and Technology is investigating human-robot interaction from psychological perspectives (Prof. Dr. Astrid Rosenthal-von der Pütten). These lab visits and exchanges help to understand better how human-machine adaptation can be investigated and what kind of fits and non-fits in this adaptation process can be observed between humans and robots. Who adepts to whom? Based on this qualitative-empirical study, aspects of 'activity as technology' by asking the following questions: What does it conceptually mean, when technology developers try to set up complex technical assemblage for operational-

izing and materializing complex agency? How is human-machine adaptivity ensured in case of machinic actions? What are the pitfalls for aligning human and machinic actions? — could be clarified.

The qualitative-empirical study allows to identify the main machinic actions in adaptive robotics. The literature study provides a list of various actions and their operationalization for adaptive robotics. However, the philosophical task — to classify these actions according to conceptual criteria — remains. These criteria can be types, levels, and scales of machinic actions.

— Types of machinic actions: A first look at the literature exhibits various types of actions such as machinic perceiving, machinic moving, machinic classifying, machinic pattern recognition, etc.

— Levels of machinic actions: Similar to Mitcham's 'technology as activity' approach basic and multiplex actions have to be distinguished. For instance, machinic learning and machinic adapting are composed of various forms of machinic actions. In fact, these multiplex actions will be cornerstones of a philosophy of machine agency. If machinic perception, cognition, and decision are multiplex actions, they have to be analysed technically as well as philosophically.

— Scale of description of machinic actions: Analysing the various levels and modes of machinic actions call for some decisions in choosing the appropriate scale of description of machinic actions. What is the right scale for describing machinic actions and machine agency, respectively? Is the analysis of software code and ML methods required?

Analysing the various types and levels of machinic actions and setting the right scale of description will provide the basis for the investigation. The empirical study will unveil, if more classification criteria beside the above-mentioned ones will be useful for a philosophy of machinic actions and machine agency.

Having a solid basis of empirically inspired machinic actions and an initial classification and depth of description of these actions, the next step will be to change the angle of view from technology descriptions to philosophy of action and agency. The main task is to develop the conceptual elements of a philosophical theory that allows for better understanding of 'activity as technology.' The aim is to derive at a matrix of conceptual elements. Some of these elements can already be mentioned, others have still to be developed:

— Differentiation between machinic activities, multiplex actions and complex machinic actions: Most important is the specification of machine actions by distinguishing simple activities from multiplex and complex actions. While technology in general exhibits various forms of activities, for instance mechanical machines carrying out mechanical activities (e.g., steam engines carry out movements), advanced technologies carry out more complex actions. How are complex actions composed? Are complex actions multiplex actions and vice versa? What do complex actions add to machine agency? Is, for instance, adaptability only possible for machines capable of complex actions?

— Modes of machinic actions: Beside types and levels, machinic actions will exhibit different modes. Such modes can be technical modalities such as necessities and

possibilities, statistical and stochastic operations of actions, etc., but can also be more philosophical ones. In my paper on "Performativity of Actions" (Laktionova, 2022) I have developed a matrix of modes involved in the performance of human actions. For instance, performativity of human actions involves processing as metaphysical mode, knowing as epistemological mode, and acknowledging as normative mode of evaluation. It is interesting to check these types and levels on machinic actions, if and how these modes can be applied or transformed for machine agency, or, if different modes are required. The main methodology applied here is to compare concepts from philosophy of actions and agency with the outlined specifications for complex machinic actions and machine agency.

To outline an agency-based philosophy of (advanced) technologies, my guiding hypothesis for arriving at a concise agency-based philosophy — is that an agency-based philosophy of (advanced) technologies entails not only a philosophy of machinic action, but of aligned man-machine interactivity. In this respect, adaptivity is regarded as an accomplished attempt to (inter-)act. However, this alignment between humans and machines can go wrong. In order to better analyse this fits and non-fits of man-machine alignment from a philosophical perspective, several aspects are of interest:

— What forms of actions align with humans? Ought they necessarily be complex actions?

— Which role do (mis-)interpretations of the chain of actions (from both sides) play?

— Can man-machine interactivity be regarded as shared action or as collective action?

— Would, in particular, collective actions, according to John Searle (1990), require "we-intentions"? If so, do minimal machinic "we-intentions" exist (e.g., programmed into machinic behaviour or retrieved from machine learning as behavioural patterns) and are such minimal machinic "we-intentions" sufficient to explain collective actions between machines and humans?

— Is man-machine interactivity a subtle "meshing of sub-plans" (Bratman, 1992, 2007), in which machines relay on explicit sub-plans and situational scripts or learned patterns?

— How much "acknowledging" (Cavell, 1976) is required from humans for machines in order to make man-machine interactivity successful?

As technology is the result of ages of anthropological creation of artefacts, but also of aligning artefacts to human purposes, something has been turned around in case of 'activity as technology.' An agency-based philosophy of advanced technology attempts to grasp this shift.

— Do 'activity as technology' make technical artefacts like adaptive robots behave 'lifelike'?

— Becomes this lifelikeness increasingly a part of our lifeworld?

— Does the increasing machine ability to adapt force humans to adapt to machines too?

— Does this mean that advanced technologies are increasingly established as agents of our "common world", increasingly formed by hybrid we-intentions?

— How is Wittgenstein's "other-mind scepticism" (Wittgenstein, 1953; Laktionova, 2020), Cavell's "uncanniness" of shared agency (Cavell, 1988), and media theory's "uncanny valley" of avatars and robots (Mori et al., 2012) related to this "common world" agency and hybrid we-intentions of advanced technologies?

These are some aspects of an agency-based philosophy of (advanced) technologies directed toward better understanding of the shift to activity as technology. The proposed research suggestions could allow to analyse the paradigmatic case of adaptive robots in depth and relate philosophical account on empirical findings.

REFERENCES

Austin, J.L. (1962). *How to Do Things with Words*. Cambridge, MA: Clarendon Press.

Bachelard, G. (1985). *The New Scientific Spirit* (originally published in 1935, transl. by A. Goldhammer). Boston, MA: Beacon Press.

Barandiaran, X.E., Di Paolo, E., Rohde, M. (2009). Defining Agency: Individuality, Normativity, Asymmetry, and Spatio-Temporality in Action. *Adaptive Behavior,* 17(5), 367—386.

Bostrom, N. (2013). *Superintelligence. Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Bratman, M.E. (1992). Shared Cooperative Activity. *The Philosophical Review*, 101, 327—341.

Bratman, M.E. (2007). *Structures of Agency: Essays*. Oxford: Oxford University Press.

Brooks, R. (1990). Elephants don't play chess. *Robotics and Autonomous Systems*, 6(1—2), 3—15.

Capurro, R, & Nagenborg, M. (Eds.). (2009). *Ethics and Robotics*. Heidelberg: Akademische Verlagsgesellschaft.

Cavell, S. (1976). Knowing and acknowleding. In: *Must We Mean What We Say? A Book of Essays* (ed. by S. Cavell) (pp. 220—245). Cambridge: Cambridge Rafael University Press.

Cavell, S. (1988). The Uncanniness of the Ordinary. In: *In Quest of the Ordinary* (ed. by S. Cavell) (pp. 153—180). Chicago: Chicago University Press.

Dessauer, F. (1927). *Philosophie der Technik. Das Problem der Realisierung*. Bonn: Cohen.

Dessauer, F. (1972). Philosophy in Its Proper Sphere (partial translation of Philosophie der Technik). In: *Philosophy and Technology* (ed. by C. Mitcham. R. Mackey) (pp. 317-334). New York: Free Press.

Davidson, D. (1963). Actions, Reasons, and Causes. *The Journal of Philosophy*, LX, 685—700.

Dennett, D. (1971). Intentional systems. *Journal of Philosophy*, 68, 87—106.

Dennett, D. (1984). *The Intentional Stance*. Cambridge, MA: Bradford Books/MIT Press.

Dias, P. (2019). *Philosophy for Engineering: Practice, Context, Ethics, Models, Failure*. Singapore: Springer.

Dignum, V., Baldoni, M., Baroglio, C. et al. (2018). Ethics by Design: necessity or curse? In: *AIES'18: Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 60—66). S.l.

Ellul, J. (1964). *The technological society* (originally published in 1954). New York: Knopf.

EU. (2013). *Responsible Research and Innovation (RRI), Science and Technology. Special Eurobarometer 401*. Brussels: European Commission. Retrieved from: https://data.europa.eu/data/datasets/s1096_79_2_401?locale=en

EU High-Level Expert Group on AI. (2019). *Ethics guidelines for trustworthy AI*. Brussels: European Commission. Retrieved from: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419

Floridi, L. (2013). *The ethics of information*. Oxford: Oxford University Press.

Floridi, L. & Jack, W.S. (2002). Computer Ethics: Mapping the Foundationalist Debate. *Ethics and Information Technology*, 4(1), 1—9.
Frankfurt, H. (1971). Freedom of the Will and the Concept of a Person. *Journal of Philosophy*, 68(1), 5—20.
Frankfurt, H. (1978). The Problem of Action. *American Philosophical Quarterly*, 15(2), 157—162.
Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Oxford University Press.
Gallagher, S. (2008). Inference or Interaction: Social Cognition without Precursors. *Philosophical Explorations*, 11(3), 163—74.
Gallagher, S. (2020). *Action and Interaction*. Oxford: University Press.
Goldman, A.I. (1970). *A Theory of Human Action*. Englewood Cliffs, NJ: Prentice-Hall.
Gramelsberger, G. (2023). *Philosophie des Digitalen zur Einführung*. Hamburg: Junius.
Heidegger, M. (1993). The Question Concerning Technology, (originally published in 1953). In: *Basic Writings: Second Edition, Revised and Expanded* (ed. by D.F. Krell) (pp. 307-342). New York: Harper Collins.
Hornsby, J. (1980). *Actions*. London: Routledge & Kegan Paul.
Hornsby, J. (2004). Agency and Actions. In: *Agency and Action* (ed. by J. Hyman, H. Steward) (pp. 1—23). Cambridge: Cambridge University Press.
Hornsby, J. (2010). Trying to Act. In: *A Companion to the Philosophy of Action* (ed. by T. O'Connor, C. Sandis) (pp. 18—25). Hoboken, NJ: Wiley-Blackwell.
Hornsby, J. (2012). Actions and Activity. *Philosophical Issues*, 22(1), 233—245.
IEEE. (2021). *7000-2021 — IEEE Standard Model Process for Addressing Ethical Concerns during System Design (INSPEC Accession Number: 21566827)*. Piscataway Township, NJ: Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/IEEESTD.2021.9536679
Kaminski, A. (2020). Gründe geben. Maschinelles Lernen als Problem der Moralfähigkeit von Entscheidungen. In: *Ethische Herausforderungen von Big-Data* (ed. by K. Wiegerling, M. Nerurkar, C. Wadephul) (pp. 151—174). Bielefeld: Transcript.
Kapp, E. (2018). *Elements of a Philosophy of Technology: On the Evolutionary History of Culture* (originally published in 1877). Minneapolis, MN: University of Minnesota Press.
Karafyllis, N.C. (2007). Growth of Biofacts: the real thing or metaphor? In: *Tensions and Convergences. Technological and Aesthetic (Trans)Formations of Society* (ed. by R. Heil, A. Kaminski, M. Stippack et al.) (pp. 141—152). Bielefeld.
Kasprowicz, D. (2019). *Der Körper auf Tauchstation. Zu einer Wissensgeschichte der Immersion*. Baden Baden: Nomos.
Laktionova, A. (2015). Collective and Individual Actions. [In Ukrainian]. *Filosofiya Osvity / Philosophy of Education*, 17(2), 108—118. Retrieved from: https://philosopheducation.com/index.php/philed/article/view/508 [=Лактіонова 2015]
Laktionova, A. (2016). Philosophy of Action. [In Ukrainian]. Kyiv: Taras Shevchenko National University of Kyiv. [=Лактіонова 2016]
Laktionova, A. (2016a). Meta-Ontological Discourse on Being and Activity. [In Ukrainian]. *Bulletin of Taras Shevchenko National University of Kyiv, Political science*, 120, 28—31. [=Лактіонова 2016a]
Laktionova, A. (2016b). Contemporary Philosophy of Action (Practical Philosophy) in Analytic Tradition. [In Ukrainian]. *Bulletin of Taras Shevchenko National University of Kyiv, Philosophy*, 1(23), 49—54. [=Лактіонова 2016b]
Laktionova, A. (2020). Wittgenstein's method and contemporary studies of the history of analytical philosophy. (Book review of Oskari Kuusela, Wittgenstein on Logic as the Method of Philosophy: Re-examining the Roots and Development of Analytic Philosophy, Oxford University Press, 2019). *Sententiae*, 39(1), 176—184.

Laktionova, A. (2022). Performativity of Actions: Possible Mappings with Jennifer Hornsby's Views on Acting, Actions, Activity and Agency. *Bulletin of Taras Shevchenko National University of Kyiv. Philosophy*, 1(6), 40—43. https://doi.org/10.17721/2523-4064.2022/6-7/13

Laktionova, A. (2023). Philosophy of Engineering and Design (Technological) Actions. *Philosophical Thought*, 1, 148—161. https://doi.org/10.15407/fd2023.01.148

Laktionova, A. (2023a). Thematic "Infrastructure" of Philosophy of Engineering. In: *The Days of Science of the Faculty of Philosophy — 2023* (ed. by A. Konverskyi) (pp. 109—112). Kyiv: Taras Shevchenko National University of Kyiv.

Laktionova, A. (2024). Common Agency and Actions. In: International Scientific Conference "The Days of Science of the Faculty of Philosophy — 2024". [Abstracts] (ed. board: A. Konverskyi et al.) (pp. 92—95). Kyiv: Kyiv University.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.

Law, J. (1992). Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity. *Systems Practice*, 5(4), 379—393.

Leonelli, S. (2016). Locating ethics in data science: responsibility and accountability in global and distributed knowledge production systems. *Philosophical Transactions of the Royal Society,* A 374, 20160122. https://doi.org/10.1098/rsta.2016.0122.

Misselhorn, C. (2019). *Grundfragen der Maschinenethik*. Stuttgart: Reclam.

Mitcham, C. (1994). *Thinking Through Technology: The Path Between Engineering and Philosophy*. Chicago: University of Chicago Press.

Mitcham, C. (2019). *Steps toward a Philosophy of Engineering: Historico-Philosophical and Critical Essays*. Lanham, MD: Rowman & Littlefield International.

Mori, M., MacDorman, K.F., & Norri Kageki. (2012). The Uncanny Valley. *IEEE Robotics & Automation Magazine*, 19, 98—100.

Mumford, L. (1967—70). *The Myth of the Machine* (2 vols.). San Diego, CA: Harcourt.

Nolfi, S. (2022). Progress and challenges in adaptive robotics. *Frontiers in Robotics and AI 9*, 1020462. https://doi.org/10.3389/frobt.2022.1020462

Nordmann, A. (2006). Collapse of Distance: Epistemic Strategies of Science and Technoscience. *Danish Yearbook of Philosophy*, 41, 7—34.

Nyholm, S. (2023). Ethics of Artificial Intelligence. In: *Encyclopedia of the Philosophy of Law and Social Philosophy* (ed. by M. Sellers, S. Kirste) (pp. 1—8). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-6730-0_1093-1

Pantsar, M. (2023). Developing Artificial Human-Like Arithmetical Intelligence (and Why) (pp. 1—18). In: *Minds and Machines*. [In print.]

Roth, A.S. (2017). Shared Agency. In: The Stanford Encyclopedia of Philosophy (ed. by E.N. Zalta). Retrieved from: https://plato.stanford.edu/archives/sum2017/entries/shared-agency/

Schlosser, M. (2018). Embodied Cognition and Temporally Extended Agency. *Synthese*, 195(5), 2089—2112.

Schlosser, M. (2019). Agency. In: The Stanford Encyclopedia of Philosophy, (ed. by E.N. Zalta). Retrieved from: https://plato.stanford.edu/archives/win2019/entries/agency/

Schomberg, R. von. (2013). *A Vision of Responsible Research and Innovation. In: Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (ed. by R. Owen, J. Bessant) (pp. 51—74). London: John Wiley.

Searl, J. (1990). Collective Intentions and Actions. In: *Intentions in Communication* (ed. by P. Cohen, J. Morgan, M. Pollack) (pp. 401—415). Cambridge: MIT Press.

Shapiro, L. (2007). The Embodied Cognition Research Programme. *Philosophy Compass*, 2(2), 338—346.

Shapiro, L. (2013). Dynamics and Cognition. *Minds and Machines*, 23(3), 353—75.

Shapiro, L. (2019). *Embodied Cognition*. London, New York: Routledge.
Wang, S. (2019). Why Adaptive Robots are the Next Big Thing. In: *roboticsbusinessreview.com*, 29.7.2019. Retrieved from: https://www.roboticsbusinessreview.com/content-from-our-sponsor/why-adaptive-robots-are-the-next-big-thing/
Wittgenstein, L. (1953). *Philosophical Investigations* (ed. by G.E.M. Anscombe, R. Rhees, transl. by G.E.M. Anscombe). Oxford: Blackwell.

Received 28.10.2024

Анна ЛАКТІОНОВА,
доктор філософських наук, професор, кафедра філософії науки та технології, Рейнсько-Вестфальський технічний університет, Аахен (Німеччина);
професор кафедри теоретичної та практичної філософії філософського факультету, Київський національний університет імені Тараса Шевченка, 01601, м. Київ (Україна), вул. Володимирська, 60
laktionovaanna@yahoo.com
https://orcid.org/0000-0001-8236-7217

ВЗАЄМОДІЯ ЛЮДИНИ І МАШИНИ:
ВИРІВНЮВАННЯ, АДАПТАЦІЯ, БУТТЯ ДІЄВЦЕМ

У статті окреслено головні моменти проєкту «Зрушення до ґрунтованої на здатності до дій філософії (передових) технологій». Основний сюжет цієї розвідки стосується взаємодій між людьми та машинами й відповідної інтерпретації взаємного вирівнювання, адаптації залучених до таких взаємодій дієвців, а також статус дієвця як такий. Щодо теоретичних та історичних підстав проєкту можна згадати такі дослідницькі галузі, як філософія науки, філософія технології, філософія інженерно-конструкторських технологічних дій, STS (соціальні дослідження науки і технологій), прикладна етика тощо. Розглянути активність (здатність до дій) як технологію, переосмислити роль активності (здатності до дій) у технології є найамбітнішою метою зазначеного проєкту: «активність (здатність до дій) як технологія» зумовлює фокусування на активності (здатності до дій) завдяки технології. Англомовні терміни «agency» та «activity» використовуються в цій статті як синоніми з огляду на основне аристотелівське значення активності (здатності до дій) як потенційної здатності дієвця до дій. Запропонована автором теорія дії та активності (здатності до дій) (і, відповідно, захищена у 2016 році дисертація доктора філософських наук) застосовується для філософських переосмислень різноманітних проблем, що стосуються поточного неперервного вдосконалення сфер штучного інтелекту; машинного навчання (включно з глибинними методами навчання та методами симулювання); винайдення і вдосконалення комп'ютерів та кодів; машин та роботів з числовим керуванням; пристроїв, оснащених комп'ютерними мікросхемами; «розумних об'єктів» тощо.
Отримати емпіричні матеріали для описаного у статті дослідницького проєкту про взаємодії між людьми та машинами можна від співпраці, наприклад, із такими «лабораторіями» німецького Аахенського університету RWTH: Інститут промислової інженерії та ергономіки (https://www.iaw.rwth-aachen.de/cms/~ieplw/IAW/lidx/1/); Конструкторський центр робототехніки (https://construction-robotics.de/en/); Кафедра особистості та технології (https://www.itec.rwth-aachen.de/cms/~sjbp/ITEC/?lidx=1).

Ключові слова: *активність (здатність до дій), (взаємо)дії, дієвець, машина, вирівнювання, адаптування, штучний інтелект.*